Notes from a site visit to *Zusha!* and *gui2de*'s program in Nairobi, Kenya, February 2-3, 2017

GiveWell staff on this site visit

- Josh Rosenberg – Senior Research Analyst, GiveWell
- Chelsea Tabart – Research Analyst, GiveWell

Zusha! or *gui2de* staff

- Dr. William Jack – Professor of Economics, Georgetown University
- Whitney Tate – Program Director, Georgetown University Initiative on Innovation, Development and Evaluation (*gui2de)*
- Elaine Hurt – *Zusha!* Program Director
- Josiah Muyesu – Senior Project Assistant, *Zusha!*
- Mallory Holding – Project Manager, *gui2de*
- Pauline Kiconco – Project Assistant, *gui2de*
- Brian Corry – Project Manager, *gui2de*
- Stella Bororo – Project Assistant, *gui2de*

Note: These notes were compiled as an informal summary of the site visit by GiveWell and primarily provide Josh's perspective on the visit. The notes originated as an internal report to other GiveWell staff.

A note on terminology in these notes: In these notes, we generally refer to the program we're interested in as "Zusha!" to distinguish it from Georgetown University Initiative on Innovation, Development and Evaluation (gui2de)'s other programs (not related to road safety). However, Zusha! is only the name of the road safety campaign in Kenya; the road safety campaigns in Rwanda, Tanzania, and Uganda have other names. We've used the term Zusha! in these notes for simplicity.

Overview

I thought this visit was useful for at least a few reasons:

1. I hadn't yet worked hard to understand the details of how the program operates. We got to see roughly all aspects of the program in Kenya, so we have a much better understanding of that now.
2. We learned a lot more about Zusha!'s basic plans for monitoring in Kenya and what challenges it faces. I think this will set us up better to review and understand its monitoring later this year. We also discussed some new ideas for long-term follow-

ups on previous Zusha! randomized controlled trials (RCTs) that I think have the potential to be useful.
3. It was a good opportunity to ask a lot of questions about all of the RCTs: the two previously done in Kenya, and the shortly upcoming ones in Tanzania and Uganda. It helped to emphasize that the results from Tanzania and Uganda may not be fully representative of what a larger scale program in those countries would look like (more details below).

As a reminder, I think the major remaining questions that will determine whether we recommend Zusha! are:

1. What will the RCTs in Tanzania and Uganda find? (Rwanda RCT is still on hold. We still expect to get results from Tanzania in ~Sept 2017 and first half of Uganda in Sept 2017, with second half in ~March 2018.)
2. Will Zusha! have high-quality monitoring in Kenya that shows whether it is successfully reaching a large proportion of buses with stickers?
3. Will Zusha! have high-quality monitoring that can help us determine whether the intervention remains effective over time in Kenya?

Some details we learned about how the program operates

Prior to this trip, I didn't have a great sense of what the actual sticker distribution process in Kenya looked like and who was doing what. Now I have a much better sense.

Zusha! estimates they've reached ~33K buses in Kenya at this point. The basic process as I understand it now is:

- They've distributed the bulk of their stickers through the insurance company they're partnering with (called "DirectLine Assurance"), which insures ~60% of the relevant vehicles in Kenya and has 17 insurance centers across Kenya. Insurance is required for all public service vehicles (PSVs) in Kenya (i.e., for all matatus and buses). Typically, an "insurance agent" - someone who is responsible for purchasing insurance for 5-100 vehicles (we couldn't get an average number, but it seemed like it might be in the 10-50 range) - will come in to the DirectLine office when the insurance for a portion of the vehicles the agent covers is near expiration (insurance is purchased for ~1 week to ~1 year at a time). The agent is usually buying insurance for 2-10 vehicles at a time (ballpark), since insurance usually expires at different times for different vehicles. DirectLine's electronic system automatically checks whether the vehicles that the agent is purchasing insurance for have been issued stickers within the last 6 months. If they haven't, or if the insurance agent just requests replacement stickers, the insurance agent will give Zusha! stickers (4 stickers for matatus, 8 stickers for buses (larger vehicles)) and explain the program.

The sticker packets (which we got samples of) also have an explanation of the lottery on them.

- Before learning all of this, I hadn't appreciated that the chain between sticker distribution and putting the stickers in buses is pretty long. Basically, the chain is often something like: Zusha! staff -> DirectLine staff -> insurance agent -> vehicle owner or "Sacco" (basically vehicle cooperative) representative -> driver. So, the stickers and the explanation of the program have to be passed between many people. This makes it seem much more likely that the stickers get lost somewhere in the chain or that the program is explained poorly to drivers.
- Every week, DirectLine and Zusha! (under license and observation of the Kenya Betting Control and Licensing Board (BCLB)) jointly run a lottery to pick winning vehicles that have received stickers. They use DirectLine's records of which vehicles have been issued stickers to choose winners. In order to get the lottery payout, the vehicle needs to actually have the stickers displayed in their vehicle. An insurance claims investigator from DirectLine checks whether the vehicle has stickers displayed as part of his or her general investigation work. If a vehicle wins and displayed the stickers, ~$50 is paid to the vehicle driver, ~$50 to the vehicle owner, and ~$50 to the insurance agent for that vehicle. DirectLine pays the cost of the lottery, arranges for payments to the winners via mobile money, handles the compliance aspect of those payments, and is a strong supporter of the program. The lottery system ensures that winners are geographically distributed across Kenya.
- BCLB's issue of a license every three months is critical to Zusha! able to continue to run its lottery. Difficulties in gaining a license (which process involves various requirements, including holding the prize money in a separate bank account) meant that the lotteries were not running when we visited Kenya.
- Zusha! has also more recently begun distributing stickers through the Kenya National Transport and Safety Authority (NTSA). The NTSA has 17 inspection centers all over Kenya. All PSVs are required to receive an inspection from NTSA at least once per year. NTSA inspects vehicles for road-readiness, safety, etc. Zusha! is now closely partnered with NTSA. When a vehicle shows up to an NTSA inspection center, a staff member at the entry point is supposed to ask the driver whether they have Zusha! stickers. If they don't, the staff member is supposed to issue stickers and explain the program. Then, when the vehicle is actually being reviewed by an NTSA inspector, the inspector is supposed to check whether the stickers have been displayed and, if not, is supposed to put up the stickers for the driver. So, theoretically all relevant vehicles in Kenya should receive Zusha! stickers through this process if they don't have them already. It's worth noting that Zusha! isn't currently running a lottery for vehicles that receive stickers from NTSA (since DirectLine pays lottery winners, and many of the vehicles reached by NTSA aren't insured by DirectLine).

Worth noting: the DirectLine sticker distribution process described above is, I believe (based on conversations with Zusha! staff), the same process they used for delivering stickers in the second RCT. So it at least doesn't seem to be a major external validity concern. However, it seems like a more indirect process than what they're doing in Tanzania and Uganda (where my impression is their staff/enumerators are directly putting up the stickers in treatment buses themselves and directly explaining the program to drivers) and than what they did in the first Kenya RCT (where I'm pretty sure they also put up the stickers themselves).

We don't have a sense of how scale-up in Tanzania and Uganda would be implemented and those scale-ups have the potential to be more costly than Kenya; Zusha! isn't working with insurance companies in those countries at this point and I think do not yet have partnerships with, e.g., government inspection centers.

Zusha!'s plans for short-term monitoring in Kenya

Zusha! knows that monitoring is one of our most important open questions, so we spent quite a while talking with them about their plans. I'd separate the goals of monitoring into: 1) understanding whether stickers are being used in a high proportion of targeted buses ("short-term"), 2) understanding whether the impact of stickers in buses persists over time ("long-term"). Focusing just on #1 here, Zusha! has plans to do 3 types of monitoring:

1. Monitoring at NTSA inspection centers. As discussed above, Zusha! has a close partnership with NTSA and recently began distributing stickers at all of their inspection centers. Zusha! also asked NTSA to track whether vehicles are using Zusha! stickers, to help them get a sense of their penetration (if this process were working well, Zusha! would theoretically check whether all relevant vehicles in Kenya had stickers each year). Zusha! told us that they were still working on the process for monitoring at NTSA centers, and what we saw on the site visit suggested this process still needs work. In particular, a few issues that we saw were: 1) inspectors are supposed to record whether stickers are displayed, but in the current process they aren't able to distinguish between whether stickers are displayed because the driver received them minutes ago when they arrived at the inspection center, or whether they were displayed before arriving at the inspection center, 2) at the inspection center we visited, it seemed like records of whether stickers were being used weren't being consistently kept (we asked to see the records and they didn't have them; one inspector seemed to be keeping some records but wasn't using the Zusha!-designed form, which could complicate data collection), 3) at the inspection center we visited, a staff member at the entrance is supposed to distribute stickers and explain the program; but, when Chelsea asked a driver we

saw about whether the staff member did this, the driver said that he received stickers but had no idea what the program was and didn't plan to put them up.

Going forward, Zusha! has at least a couple of options for trying to improve its monitoring at these centers: 1) It's considering putting a full-time Zusha!-paid staff member at every NTSA center to distribute stickers and document whether vehicles were using stickers. This could improve quality of distribution and monitoring since they wouldn't have to rely on busy NTSA people to add this to their workstream. 2) It's working with the NTSA and it may be able to get its monitoring questions added to NTSA's electronic system which records the output of all vehicle inspections (this would make capturing the data much easier for all parties; currently Zusha! monitoring requires a separate workstream).

2. Monitoring at bus parks (i.e., locations where many PSVs pick up passengers and begin their routes). Zusha! plans to hire staff to conduct random checks of vehicles at bus parks across Kenya, beginning in March. They haven't done this before and are still in the middle of designing the methodology and process, but basically they'd design a process to choose bus parks randomly, and vehicles within those parks randomly, and see whether vehicles have stickers (and possibly collect other information). Most of the details of this are still being worked out, so we didn't dig into it much.
3. Monitoring of whether buses are using stickers via the lottery. Zusha! told us that they can't currently use the information from lottery inspections as evidence of sticker usage because they believe that drivers are "gaming" the system. As it works now, DirectLine calls the vehicle driver to let them know they've been selected in the lottery and that they're coming for a sticker inspection. So, the driver may not have been using the stickers but could then put them up or otherwise find Zusha! stickers to put up pre-inspection. They think drivers are gaming the system because they noticed that some of the vehicles that succeeded in the lottery weren't using stickers according to NTSA monitoring. Zusha! has some ideas for how they might improve this (e.g., maybe making the back of the sticker a lottery ticket that needs to be submitted for lottery eligibility), but are still in the process of deciding how and whether to change the system. As it stands, if monitoring processes #1 and #2 above were working well, they may not need to solve this issue. Instead, the lottery could primarily function as a way to incentivize sticker use (which they see as its primary purpose now).

So, basically, all of Zusha!'s current monitoring processes have issues or aren't well-defined, but they know this and are in the middle of trying to significantly improve their processes. They expect to have new processes in place in March and to share data from the first few months of their new monitoring in ~June. It's unclear how successful they'll be,

but it was heartening to see that they were being pretty critical about their monitoring processes and being proactive about ways to improve.

We also asked why monitoring at bus parks would be useful if they are able to do monitoring at NTSA inspection centers well. They told us that they want to triangulate information and have more frequent checks; they think checking vehicles only once per year at NTSA centers may not give a great sense of whether stickers are actually consistently being used in buses. One aspect of this is that they are working on creating a database that collates their three streams of monitoring and allows compliance to be compared across streams.

Zusha!'s plans for long-term monitoring in Kenya

A major question about the impact of Zusha!'s intervention is whether it persists over the longer term. Since it's a behavioral intervention, it's easy to imagine the impact weakening over time, and one can also imagine it strengthening over time if Zusha! becomes culturally accepted. I don't think we're going to be able to get very good information about whether either of these are occurring, but we still wanted to think about the best we could do. We discussed some interesting new opportunities. In brief, some ideas are:

1. Potential "accidental experiment." In Zusha!'s second RCT, they used DirectLine's electronic systems to allocate ~20% of vehicles to the control group. DirectLine recently told them that they thought they might've accidentally forgotten to turn off the algorithm that stopped a random ~20% of buses from receiving stickers. If that's actually the case (DirectLine and Zusha! are still investigating), then Zusha! may have another 'accidental' RCT evaluation that they can do of their program in Kenya. I.e., they could see what accident rates were like in the accidentally-control buses vs. treatment buses.
2. Long-term follow-ups to previous RCTs. Zusha! has a lot of data on accident rates in DirectLine-insured buses going back to 2008. In their first RCT, they stopped comparing treatment vs. control buses after one year (~in 2010) because they thought that was a reasonable amount of time to check the impact of the intervention. In the second RCT, they did something similar. But, it seems that they may be able to do a longer term historical comparison of treatment vs. control buses across the two RCTs. It won't be a perfect experiment since in many cases the control buses will have been reached by stickers following the RCTs, but it could still be informative to see whether the accident rates converged over time or whether the treatment buses persisted in having lower accident rates. Also, it may be that many buses were allocated to the control group in both RCTs, which would provide another interesting random comparison. Dr. Jack said that they would consider

exploring this idea. It could give us some high-quality information about the possible long-term effects of the program relatively soon.

3. Looking at accident rates in Kenya overall and accident rates in buses/matatus versus other types of vehicles. Zusha! told us that they have aggregate data on accident rates over time in DirectLine vehicles and possibly in other types of vehicles. This won't be randomized, but can give us a general sense of accident trends over time and could be interesting if buses/matatus improved a lot relative to other types of vehicles. Zusha! told us they'd share what they have on this soon. This would also be useful for thinking about whether we should expect the reduction in deaths to be much lower now, since accident rates in general in Kenya may have improved a lot since the time the RCTs were done (second RCT was 2011-2013 I think).

What we learned about the RCTs

New RCTs

- The Tanzania RCT is only being done in large buses (25+ passengers) that travel long distances (drivers are asked if there is the potential to go >80 km/hr on their trip). They did this because those vehicles have the most potential for deadly accidents, and also enable them to get as much power for the study as possible. (This also makes finding significant effects in this study slightly more likely than for the 2nd Kenya RCT, which included short-distance and low-speed buses.) But that means that if the Tanzania RCT finds strong positive results, there's pretty limited scale-up potential within Tanzania. They're already reaching ~all of the large buses they could find through the RCT (~2,000 buses control, ~2,000 buses treatment, very roughly). So, if they find positive effects they can deliver stickers to these 4,000 buses going forward, but it's not clear you'd want to scale up to smaller vehicles (since the RCT didn't test if it was effective in those types of vehicles). I think it'd probably make sense to try to reach some smaller vehicles, but we'd want to discount the cost-effectiveness relative to the RCT results. This is a significant concern about funding for scale-up that hadn't been as salient in my mind before the trip.
- The Uganda RCT started off being done in only relatively large-passenger vehicles that travel long distances (similar sample size to Tanzania). But, because of our grant, Zusha! now hopes to also reach ~2,000 buses that travel at lower speeds and shorter distances. This seems like a pretty useful impact of our grant since it could potentially give us a better sense of how effective the stickers are in a population of buses where the program could be scaled more, but the sample is still quite limited.

- Collecting data on accidents in Tanzania has been a challenge. The police have been hesitant to let Zusha! view accident data directly. Instead, the police agreed to fill out a special form for Zusha!. A major question is whether the police are reporting all accidents; if they're not, it could reduce the power of the study. There's precedent for incomplete reporting: the police have an electronic accident recording system that is rarely used. At first we thought this was likely to greatly reduce the power of the study. But after talking with Dr. Jack, I feel less worried about this; it seems like the number of accidents they've collected is in line with the assumptions in their power calculation (~10% of buses get in an accident each year). There's also a concern about the quality of the accident data. Zusha! ran a validation process where they checked 20 of the accidents that were reported (out of hundreds to thousands) to see if they matched police records; this was apparently a cumbersome process so they couldn't do more. 16 of 20 records matched. The checks were on recorded accidents from a range of police stations, sampling one or two records from each. The data quality is still a concern, but seems like it should add noise (thus reducing power) rather than systematically biasing the study. As far as we and Zusha! know, the police are blinded to the treatment status of buses.

Kenya RCTs

- We talked with Dr. Jack a bit about the fact that placebo stickers found similar effects to treatment stickers (though placebo was barely statistically insignificant while treatment was barely statistically significant). Dr. Jack said that he agreed the mechanism of how the intervention is having an impact is still unclear to him. However, he and Zusha! for now are content with the intention-to-treat estimates from the studies which show impact even though they do not necessarily identify mechanisms, especially since a study designed to identify mechanisms would likely be very costly and may not be feasible. He also said that he thought the placebo was imperfect (stickers that said things like "Travel well" could still conceivably have an impact; they wanted the stickers to say things like "Drink more milk," but it would have been difficult for DirectLine to explain to recipients why they were promoting milk).
- We asked Dr. Jack a question about imbalance between historical claims in the control versus treatment groups. He said that the randomization process seemed solid, buses seemed balanced on most characteristics, and that the way buses were imbalanced suggested that the treatment effect would be underestimated. He thinks that there wasn't enough imbalance to suggest significant flaws in randomization. I don't understand the stats here as well as I would like, so this is something we could consider digging into further.

Other general methodological concerns:

- Drs. Jack and Habyarimana are doing the analysis of the RCT results, but they are not independent evaluators since they also run the Zusha! program. My impression is that RCTs are often done by external evaluators to avoid this kind of issue. Dr. Jack said that he understood this is a concern and would be interested in sharing the data set with other parties so that they can replicate and analyze the results.
- Another concern is that many of the people involved with the program/RCT aren't blinded to the treatment status of vehicles. In the Uganda and Tanzania RCTs, license plate numbers that end in an even number are treatment, and odd are control (or vice versa; I don't remember). We couldn't come up with a clear way that this would affect the results, but seems worth flagging and having in mind.

Other miscellaneous content

Cost per vehicle reached

- We asked a lot of questions to get at whether major costs of the program were being paid by other actors, and whether we were missing any indirect costs.
    - DirectLine (the insurance company) pays lottery costs, and I feel pretty confident we should exclude this from our CEA (since DirectLine is a private company, I don't think opportunity cost of those dollars is high). We previously included this in our CEA, so this would make a very slight improvement (this was ~5% of all costs or something like that). It's worth noting that this may change in the future, as Zusha! may eventually target all buses with the lottery in Kenya, meaning that government or philanthropists may have to shoulder some or all of its cost. In addition, insurers are not yet an implementing partner in Tanzania or Uganda, meaning Zusha! or government could pay the lottery costs of those programs.
    - I'm not sure that we factored in all of Dr. Jack and Dr. Habyarimana's time to the costs of the program. I think we probably did, but something to consider in the future. Zusha! takes something like ~25% of Dr. Jack's time and ~10% of Dr. Habyarimana's time.
    - There were some other miscellaneous items that we noted down, but the most important outcome is that I think Zusha! has a strong sense of what costs we're interested in now, and told us they'd have an updated comprehensive cost estimate in September in time for our review.

RFMF

- Sounds like Zusha! still doesn't have much in the way of other potential funders, but also that it has not made fundraising outside of USAID and East African funders (i.e.,

government, DirectLine) a priority to date. Though, we did not discuss this topic in detail yet due to time constraints.

Miscellaneous general impressions of the organization/people

- We were generally impressed with them as an organization. They were very open, seemed competent and thoughtful, and were good at answering questions. Ms. Tate, Dr. Jack, and Ms. Holding stood out as being especially helpful.

Other tidbits

- Prior to our grant, Zusha! thought it was likely that they would need to pass off the program to the government as soon as possible or risk letting the program end. Now that we've made a grant, they told us they expect to pass off the programs more slowly but that they think passing off to the government is more likely to succeed. This is an interesting effect of our grant; it seems like this is probably good, but is a possible offsetting impact—i.e., perhaps we slowed the progress of the government taking over the program or paying for more of the program sooner.
- We saw a pretty good number of Zusha! stickers in vehicles. I think they were in roughly ~40% of the buses/matatus I saw around Nairobi, though I definitely wasn't doing any kind of random or systematic check.
- We asked lots of people we came across if they knew about Zusha!. On Saturday morning, we went to Nairobi National Park on personal time. Our driver said he knew about Zusha! and had Zusha!'d the previous day when his driver was driving badly. He said the stickers helped him speak up, though he didn't specify in what way, and we didn't press.
- A thought we had for monitoring was that maybe Zusha! could check people's awareness of the intervention somehow instead of just checking for stickers in buses. Zusha! said they're planning to consider something like that. (We made clear they shouldn't just do it because we asked about it.)

Next steps

As mentioned above, I think major next check-in points will be:

- June: should have a few months of results from new monitoring processes in Kenya. We'll also follow up to see how the long-term monitoring is coming along.
- September: should have results from Tanzania, partial results from Uganda, more monitoring data from Kenya, cost per treatment analysis. This will be the time we really gear up and evaluate them for top charity status.

What we did

Day 1 (~9 hours of total time with Zusha! staff)

- A few hours of meetings at gui2de's main offices where they gave us an overview of the Tanzania and Uganda RCTs and plans for monitoring in Kenya, and we asked lots of questions.
- Social lunch with gui2de staff + advisors.
- Visited a bus park, where we spoke to a driver, asked questions to Tanzania and Uganda teams about enrolling drivers in the RCTs.
- Met with Dr. Jack and Ms. Tate. Asked Dr. Jack a lot of questions about the RCTs.

Day 2 (~8 hours of total time with Zusha! staff)

- Drove to the insurance sales office with Ms. Tate; asked a lot of general questions about Zusha!.
- Visited an insurance sales office, which is where stickers are distributed to insurance agents, who then give them to owners/saccos/drivers. Asked questions about how they distribute stickers, what insurance sellers communicate to insurance buyers about Zusha!, etc.
- Learned more about lottery process at DirectLine (insurance company). Learned how randomization was done. Met investigators who check the stickers in buses. Talked about potential flaws in that process.
- Got lunch with gui2de and one DirectLine staff.
- Took a matatu to NTSA. The matatu was privately hired for us, though, so wasn't packed with people as would be typical.
- Went to NTSA inspection center. Talked to policeman at gate who gives out stickers. Talked to NTSA managers: one manager for that branch and one manager for all NTSA branches.
- Debriefed with Ms. Tate and Dr. Jack and talked about next steps.

All GiveWell site visit notes are available at http://www.givewell.org/international/site-visits